Mini Proyecto 1.

Gold Price Prediction

Context

Historically, gold had been used as a form of currency in various parts of the world including the USA. In present times, precious metals like gold are held with central banks of all countries to guarantee re-payment of foreign debts, and also to control inflation which results in reflecting the financial strength of the country. Recently, emerging world economies, such as China, Russia, and India have been big buyers of gold, whereas the USA, So USA, South Africa, and Australia are among the big seller of gold.

Forecasting rise and fall in the daily gold rates can help investors to decide when to buy (or sell) the commodity. But Gold prices are dependent on many factors such as prices of other precious metals, prices of crude oil, stock exchange performance, Bonds prices, currency exchange rates, etc.

The challenge of this project is to accurately predict the future adjusted closing price of Gold ETF across a given period of time in the future. The problem is a regression problem, because the output value which is the adjusted closing price in this project is continuous value.

Content

Data for this study is collected from November 18th 2011 to January 1st 2019 from various sources. The data has 1718 rows in total and 80 columns in total. Data for attributes, such as Oil Price, Standard and Poor's (S&P) 500 index, Dow Jones Index US Bond rates (10 years), Euro USD exchange rates, prices of precious metals Silver and Platinum and other metals such as Palladium and Rhodium, prices of US Dollar Index, Eldorado Gold Corporation and Gold Miners ETF were gathered.

The dataset has 1718 rows in total and 80 columns in total. Data for attributes, such as Oil Price, Standard and Poor's (S&P) 500 index, Dow Jones Index US Bond rates (10 years), Euro USD exchange rates, prices of precious metals Silver and Platinum and other metals such as Palladium and Rhodium, prices of US Dollar Index, Eldorado Gold Corporation and Gold Miners ETF were gathered.

The historical data of Gold ETF fetched from Yahoo finance has 7 columns, Date, Open, High, Low, Close, Adjusted Close, and Volume, the difference between Adjusted Close and Close is that the closing price of a stock is the price of that stock at the close of the trading day. Whereas the adjusted closing price takes into account factors such as dividends, stock splits, and new stock offerings to determine a value. So, **Adjusted Close is the outcome variable which is the value you have to predict.**

Features.

The historical data of Gold ETF fetched from Yahoo finance has 7 columns, Date, Open, High, Low, Close, Adjusted Close and Volume, the difference between Adjusted Close and Close is that closing price of a stock is the price of that stock at the close of the trading day. Whereas the adjusted closing

price takes into account factors such as dividends, stock splits and new stock offerings to determine a value. We would use Adjusted Close as our outcome variables which is the value we want to predict. 'SP_open', 'SP_high', 'SP_low', 'SP_close', 'SP_Ajclose', 'SP_volume' of S&P 500 Index, 'DJ_open','DJ_high', 'DJ_low', 'DJ_close', 'DJ_Ajclose', 'DJ_volume' of Dow Jones Index, 'EG_open', 'EG_high', 'EG_low', 'EG_close', 'EG_Ajclose', 'EG_volume' of Eldorado Gold Corporation (EGO), 'EU_Price','EU_open', 'EU_high', 'EU_low', 'EU_Trend' of EUR USD Exchange rate, 'OF_Price', 'OF_Open', 'OF_High', 'OF_Low', 'OF_Volume', 'OF_Trend' of Brent Crude oil Futures, 'OS_Price', 'OS_Open', 'OS_High', 'OS_Low', 'OS_Trend', of Crude Oil WTI USD, 'SF_Price', 'SF_Open', 'SF_High', 'SF_Low', 'SF_Volume', 'SF_Trend' of Silver Futures, 'USB_Price', 'USB_Open', 'USB_High','USB_Low', 'USB_Trend' of US Bond Rate data, 'PLT_Price', 'PLT_Open', 'PLT_High', 'PLT_Low','PLT_Trend' of Platinum Price, 'PLD_Price', 'PLD_Open', 'PLD_High', 'PLD_Low','PLD_Trend' of Palladium price 'RHO_PRICE' of Rhodium Prices 'USDI_Price', 'USDI_Open', 'USDI_High','USDI_Low', 'USDI_Volume', 'USDI_Trend' of US dollar Index Price, 'GDX_Open', 'GDX_High', 'GDX_Low', 'GDX_Close', 'GDX_Adj Close', 'GDX_Volume' of Gold Miners ETF, 'USO_Open','USO_High', 'USO_Low', 'USO_Close', 'USO_Adj Close', 'USO_Volume' of Oil ETF USO are used as feature engineered variables.

Elements to evaluate.

1. Report on the results and conclusions. Submit a maximum of 6 pages with the following elements (more elements can be included without exceeding 6 pages):

1.1. Database preprocessing
1.2. Preprocessing and exploratory data analysis.
1.3. Analysis and performance of the models used
1.4. Results, conclusions, inferences, future work.

2. Google Collab Notebook.
3. Oral presentation. Slide.

Deadline: March 18.

Oral presentation: March 19, 20 and 21.